

# Weather Daily Data Approximation Using Point Adaptive Ellipsoidal Neighborhood in Scattered Data Interpolation Methods

Majid Amirfakhrian<sup>a,b</sup>, Faramarz Samavati<sup>a</sup>

<sup>a</sup>*Department of Computer Science, University of Calgary, Calgary, AB T2N 1N4, Canada.*

<sup>b</sup>*Department of Mathematics, Central Tehran Branch, Islamic Azad University, Tehran, Iran.*

---

## Abstract

Many papers have applied the meshless method to approximate a function by using a set of scattered data. To use a meshless method, we need to predefine a positive real number as a radius of the local sphere or a positive integer as the number of interior points. This is while the effect of a fixed number as the radius of a local sphere or as the number of interior points could be different for different parts of a complex domain. This paper contains the construction of ellipsoidal neighborhoods for the meshless interpolation and approximation methods with local support functions and local behavior. By using these new neighborhoods, the trend of local data could be found easily.

The advantage of this method over the current methods is the use of ellipsoidal local neighborhoods which include points with the most impact on the approximation of the function. We applied these methods to the data of the daily temperature and humidity of Alberta, Canada.

**Keywords:** Moving Least Squares; Inverse Distance Weighting; Interpolation; Localization; k-Nearest Neighborhood.

**2020 MSC:** 65D05, 62J10, 62H99.

---

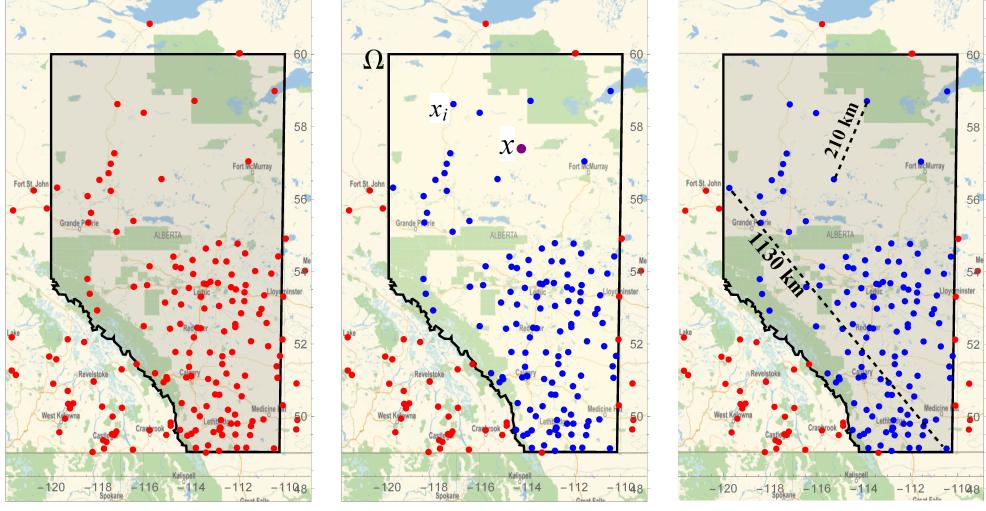
## 1. Introduction

There are many reasons that short-term and long-term weather forecasts affect our decision making. Daily weather approximations of components such as air temperature, humidity, and wind speed are important for practical situations. Generally, long term weather forecasts deal with partial differential equations that can be solved by numerical methods. Methods of solving weather forecast problems numerically typically involve divided differences, tensor products, finite elements, finite differences, and finite volumes. Another technique to solve such problems is the meshless method. Using the meshless method, the problem of interpolation may be considered over a set of scattered points without any preconditioned structure. For the climate problem,

---

\*Corresponding author

Email addresses: [majid.amirfakhrian@ucalgary.ca](mailto:majid.amirfakhrian@ucalgary.ca) (Majid Amirfakhrian), [samavati@ucalgary.ca](mailto:samavati@ucalgary.ca) (Faramarz Samavati)



(a) The location of weather stations in Alberta  
(b) The set  $\Omega$  and an arbitrary point  $x$   
(c) The maximum distance between a station and its nearest and furthest stations

Figure 1: Distribution of Weather Stations in Alberta, Canada

the structure of scattered points is the distribution of weather stations in a specific region. For our study, we  
10 consider the weather station data in Alberta, Canada. (See Figure 1(a).)

Let  $X = \{x_1, \dots, x_n\}$  be the set of the locations of all stations in the region  $\Omega \subset \mathbb{R}^2$ . Temperature, humidity and wind speed are each a function of the station location in the region, name it  $f$ . The values of function  $f$  over the set of stations are known as  $f_{i,} = f(x_i)$ , for  $i = 1, \dots, n$ . (See Figure 1(b).)

To solve the weather daily data interpolation problem, we need to approximate the values of  $f$  over all points  
15 of  $\Omega$ . So, for an arbitrary  $x \in \Omega$  we need to approximate  $f(x)$ .

The question is whether to use all points  $x_i$  in  $\Omega$  for approximating  $f(x)$  (i.e. global methods) or, alternatively, consider a subset of points in a local neighborhood of  $x$ . When a data point  $x_i$  changes, it is commonly expected to get a local modification of the approximating function; therefore, local methods are preferred over global methods. Additionally, global methods require more computations than local methods.

20 For the local method, a neighborhood is defined for each data point, and locations within the neighborhood will be used for the approximation of that data point. The challenge is the definition of a proper neighborhood for each point  $x$ . The structure of this neighborhood can affect the accuracy and the cost of computations. There are two approaches to defining a local domain. The first approach defines a disk of specific radius as the neighborhood without constraining the number of points in that neighborhood, while with the second approach,  
25 a predetermined number of points are considered as the neighborhood. Local meshless methods such as local

Radial Basis Functions (RBF) [55], Kernel-based and reproducing kernel approximation methods [20], and Moving Least Squares (MLS) [31] are examples of the first approach, while  $k$ -Nearest Neighborhood (KNN) [4, 47] and the local least squares method [10, 11] are examples of the second one.

In the context of daily weather data approximation, Inverse Distance Weighting (IDW) has been widely used. This method is a modified version of the Shepard Interpolation method [50]. IDW is a global method where closer points have a greater effect on a specific location's approximation than points farther away.

It is clear that daily weather interpolation at a specific point is inherently a local problem, so we use local MLS. To use local methods for daily weather approximation at a given  $x$  in the domain, the following issues should be considered. First, in methods such as KNN, the effects of all  $k$  nearest points of  $x$  are the same regardless of their distances to  $x$ . Second, in methods such as MLS and RBF [5], the radius of the disk neighborhood of  $x$  should be specified as a predefined value. Additionally, the RBF method is related to a parameter named “shape parameter” and the approximating function is very sensitive with respect to this parameter.

In this paper, we introduce a novel method for defining the neighborhood of points to improve weather forecast approximation accuracy without considering a predefined number of data points in its neighborhood. This neighborhood can be applied for the approximation of a function using any kind of local methods, especially for a meshless method such as MLS, that needs a local support weight function. We then find and compare several local neighborhoods by applying them to the weather data of Alberta in 2018 and 2019. We introduce three methods for constructing a local neighborhood that does not require specification of a predefined value. Specifically, we use an ellipsoidal domain rather than a disk domain to support the weight function. In an application of these methods to daily weather data, we show that the approximations of a function by using these neighborhoods have acceptable approximations based on Root Mean Squares (RMS) error and Mean Absolute Errors (MAE).

For approximating the function at any given point, we must determine what the most impactful points in its neighborhood are, how these points are distributed in that neighborhood, and the number of points there for a local interpolation. In this paper, we introduce new localization methods to address these challenges. These localizations have ellipsoidal structures rather than spherical shape, and contain the points with the most impact on computation. Such an asymmetric local neighborhood improves the approximation accuracy in comparison to other methods.

The rest of this paper has been organized as follows. Section 2 contains a compact history of related works. In Section 3, we present some basic tools in our work including IDW and MLS methods for interpolation, various localization methods, and the Mahalanobis distance. Section 4 contains our definition of two new types of elliptical quantiles based on generalized covariance and the Mahalanobis distance. Our model is described in

Section 5. In section 6, the results of applying MLS and using our new ellipsoid neighborhoods are presented.  
60 Lastly, the conclusions are given in Section 7.

## 2. Related Works

The weather forecasting problem was first introduced in 1904 by V. Bjerkness [13] as an initial value problem. In the 1950s the first weather numerical forecast was done with a computer, and since then there have been  
65 many developments in Numerical Weather Prediction (NWP). Several partial differential equations and ordinary differential equation problems applied in weather forecasting are considered in [27].

The  $k$ -nearest neighborhood (KNN) algorithm introduced in [18] is a non-parametric method. The goal of KNN is to find an optimal value of  $k$  for a specific dataset, which is related to the structure of the dataset and the given point [38]. In kernel-based methods [3, 20], a choice of the weight vector is considered, however  
70 choosing the kernel function and its bandwidth is a challenge by itself.

Inverse Distance Weighting (IDW) is one of the most used methods for interpolating climate data. IDW was used in [16] for the estimation of rainfall in Taiwan, and it is used in [56] to estimate the temperature in the US. The general IDW method uses a parameter named  $\beta$ , which tunes the effect of sample points on the final result. In [22], the authors gave an adaptive method to control the effect of this parameter in the general  
75 IDW method.

In [32], the authors proposed an efficient method by combining the local RBFs with the Shepard method for the multivariate interpolation of scattered data sets. Kernel-based approximation methods for scattered data interpolation and their stable computation are presented in [20]. The authors explained the accuracy and optimality of reproducing kernel Hilbert space methods. A radial basis function based on the partition of unity  
80 method (RBF-PUM) is proposed in [34] and it is used in localization of the computation and because of its concurrent nature, it can be programmed parallelly. A parallel algorithm is proposed through the OpenCL standard in [15].

Kriging is also a very applied method in this area [52] which uses linear regression. Kriging represents a family of generalized least-square regression algorithms. A modification of Ordinary Kriging is used for GIS  
85 in [12]. Drift kriging is a type of kriging method which was introduced in [41]. A hybrid kriging method for interpolation on high-resolution temperature maps presented in [53].

Angular distance weighting (ADW), Kriging, thin-plate splines (TPS), and Natural neighbor interpolation (NNI) are compared as different interpolation methods over daily European climate data in [24]. In [51], daily temperature from British Columbia were approximated using the IDW, gradient-plus-inverse-distance-squared (GIDS), sea-level regression Kriging (SLRK) sea-level inverse distance (SLID) and multiple linear  
90

regression (MLR) methods. The authors compared twelve interpolation methods for daily temperature.

Ozone data in Huston, Texas are analyzed in [40] using a prediction method based on Spatio-temporal Kriging. Spatio-temporal Kriging method also used for estimation of weather variables in [19]. In [48], mean daily temperature data were analyzed by using “Central Massachusetts 1971-2000” test data. Detrended Universal  
95 Kriging (DUK), Multiple Linear Regression (MLR), CoKriging, and Spline tension are presented and compared over a region in Brazil in [14]. Additionally, a bivariate stochastic model for the Spatio-temporal field of minimum and maximum temperature is presented in [28]. Three Machine Learning algorithms, Gene Expression Programming (GEP), M5 model tree (M5), and Support Vector Regression (SVR) were used to model and estimate the point temperature of a region in [45]. Three Neural network techniques, adaptive network-based fuzzy  
100 inference system (ANFIS), Auto-regressive Integrated Moving Average (ARIMA), and Genetic Algorithm (GA) were also used in [9] for prediction of the daily temperature of Dhaka, Bangladesh. A study comparing several techniques for mapping the rainfall erosivity indices over the Ebro basin (NE Spain) in [8]. Recently, a non-stationary extreme value model formulation for rainfall is considered in [49]. Hourly temperature and ozone concentration data are analyzed by a stochastic local interaction model in [25].

105 A useful tool for analyzing a set of data points is classifying them into different categories. Many works have been done on the classification of data and using a distance for finding a neighborhood of a point. A local-mean based classification algorithm was presented in [35] based on the average and largest distance of a local region. Also, in [39] the authors presented a  $k$ -means algorithm by using Mahalanobis distances between points. The effect of statistical distance is studied on the IDW interpolation method in [3]. An ellipsoid bounding region is  
110 used as a neighborhood of a point in some literature [29, 30]. Also, a study was done on fire-related weather in Alberta by comparing different interpolation methods [26]. There is some other literature containing ellipsoidal neighborhoods and adaptive nearest neighbor [42, 57].

In this paper, we define new ellipsoid neighborhoods and test the presented method on Alberta weather data for minimum and maximum temperatures.

115 **3. Preliminaries**

In this section, we present the problem of approximation in  $\mathbb{R}^d$ , and for the problem of weather forecasting in an area, we set it to  $d = 2$ .

Given a set  $X = \{x_1, \dots, x_n\} \subset \Omega$  as the set of distinct points in  $\mathbb{R}^d$  belonging to a bounded domain  $\Omega$ , and a  
120 set of  $n$  real numbers  $F = \{f_1, \dots, f_n\}$ , which are the values of the function  $f$  (e.g. temperature) over the set  $X$ , the goal is approximating the value of  $f$  over given  $x = (\gamma_1, \dots, \gamma_d)$  in  $\Omega$ .

In the following subsections, we review two approximation methods, localization and Mahalanobis distance of data, all required for describing our methods.

### 3.1. IDW and MLS

One of the most used methods for approximating the weather data is IDW. The basic idea of IDW for approximating of  $f$  on  $x$  is to use a weighted average of known  $f$  values at  $x_i$ , where the weights are inversely proportional to distance between  $x_i$  and  $x$ . More specially, the interpolating function is defined as:

$$s(x) = \begin{cases} \frac{\sum_{i=1}^n d_i^{-\beta} f_i}{\sum_{i=1}^n d_i^{-\beta}}, & x \neq x_j, \quad j = 1, \dots, n \\ f_j, & x = x_j, \end{cases} \quad (3.1)$$

where  $d_i$  is the distance between  $x$  and the data point  $x_i$ , and  $\beta$  is a real positive number that controls the impact of  $d_i$  in the weighted average. The parameter  $\beta$  is a predefined parameter where a larger value means less effect of more distant points. Also, when  $\beta$  tends to  $\infty$ , the value of the approximating function is equal to the value of the nearest point to  $x$ .

The second method we consider is Moving Least Squares (MLS). This is a powerful method to approximate  $f$  over  $\Omega$  with respect to  $X$ . This method is introduced in [31] and it is widely used in various application of Physics [1, 36, 54], Computer Graphics [2, 33, 43, 44], Economics [6], and etc.

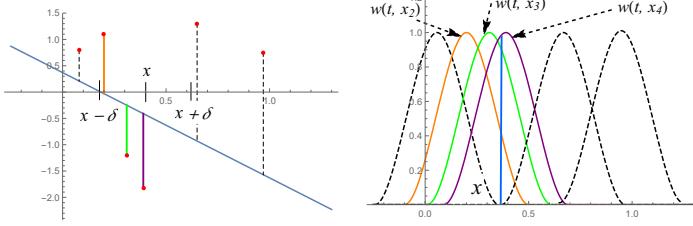
The basic idea of this method is to minimize the error of the approximating function at  $x \in \Omega$ , where the closer data points have more impacts proportional to the inverse of distance and the impact of each data point is tuned by a local weight function. Also, for any point  $x$  in the domain, a different weighted least squares problem is solved to approximate the function at  $x$ .

More formally, given a point  $x \in \Omega$ , the value of approximation is given by  $p^*(x)$  in which  $p^*(x)$  is the solution of

$$\min_{p \in \Pi_m(\mathbb{R}^d)} \sum_{i=1}^n (p(x_i) - f_i)^2 w(x, x_i), \quad (3.2)$$

where  $\Pi_m(\mathbb{R}^d)$  is the set of all  $d$ -variate polynomials of degree at most  $m$ , and  $w$  is the weight function for MLS defined by  $w(x, x_i) = \phi(\|x - x_i\|_2/\delta)$ , where  $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  is a compact supported decreasing function which its support is a subset of  $[0, 1]$ . The weight function's domain is a closed interval in 1D, a disk in 2D, a ball in 3D, and a hyper-ball in general.

Figure 2 illustrates the situation for a set of six 1-dimensional points and for a fixed polynomial  $p(x)$ . In this figure, for fixed  $x$  and  $\delta$ , three points are located in the neighborhood of  $x$  (See Figure 2(a)) and the impact of their errors are proportional to the value of corresponding weight functions at  $x$  (See Figure 2(b)). Changing the position of  $x$  affects the number of neighborhood points and the magnitude of the weight functions. In this method for every  $x \in \Omega$  we look for a polynomial with minimum error quantity at this point.



(a) Input data points sampled from an arbitrary polynomial function at  $x$   
 $p(t)$

(b) Contributing weight functions

Figure 2: MLS for point  $t$  using a set of data points in one dimension.

In MLS, not only is the number of points in the neighborhood of  $x$  important, but also the distances of them to  $x$ .

By considering a disk with radius  $\delta$ , the problem can be reduced to [7]

$$\min \sum_{i \in I_{x,X,\delta}} \left( f_i - \sum_{j=1}^q b_j p_j(x_i) \right)^2 w(x, x_i), \quad (3.3)$$

where  $I_{x,X,\delta}$  is the set of all indices of points in a disk centred in  $x$  with radius of  $\delta$ :

$$I_{x,X,\delta} = \{i \in \{1, 2, \dots, n\} : \|x - x_i\|_2 \leq \delta\}, \quad (3.4)$$

and  $\{p_j\}_{j=1}^q$  is a set of basis for  $\Pi_m(\mathbb{R}^d)$ . It is shown in [17] that  $q = \frac{m!d!}{d!}$ .

As shown in [55], to solve (3.3) it is sufficient to solve a linear system for any given  $x$ . We define

$$\begin{cases} F = (f(x_i) : i \in I_{x,X,\delta})^T, \\ P = (p_j(x_i)), \\ D = D(x) = \text{diag}(\Phi_\delta(x - x_i) : i \in I_{x,X,\delta}), \\ R = R(x) = (p_1(x), \dots, p_q(x))^T, \\ M = P^T D P, \end{cases} \quad (3.5)$$

then, we solve the system of equation

$$MG = R, \quad (3.6)$$

and finally, the approximated value is computed by  $p^*(x) = F^T D P G$ .

Figure 3 shows the best polynomials of degree 2 at three points  $x = 0.67, 0.68, 0.69$  for the data set of Figure 2, by using  $\delta = 0.3$ .

In (3.6),  $M$  is called the moment matrix of  $x$ . It can be shown that in some special cases the moment matrix is singular. A special case is given in the following remarks:

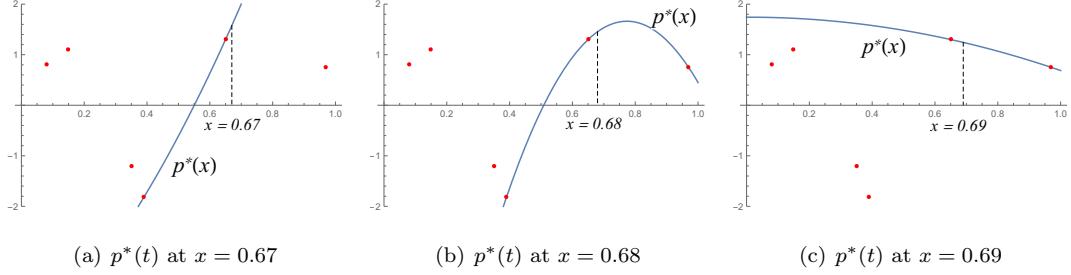


Figure 3: MLS for three close points  $x = 0.67, 0.68, 0.69$ .

**Remark 3.1.** If the number of points in the neighborhood of the given point  $x$  is less than  $q$ , then the moment matrix  $M$  will be singular.

<sup>155</sup> **Remark 3.2.** If all of the points in the neighborhood of the given point  $x$  create a subspace of  $\mathbb{R}^d$  which has less dimension than  $\mathbb{R}^d$ , then the moment matrix  $M$  will be singular.

Both Remarks 3.1 and 3.2 are consequences of this fact that the rank of matrix  $M$  is equal to the rank of matrix  $P$ . So, in this method, for each  $x$ , the number of points in the neighborhood of the given point  $x$ , must be greater than or equal to  $q$ .

<sup>160</sup> *3.2. Localization*

Localization is a vital aspect of our method. Rather than using all of the data points for determining the approximation at  $x$ , a local subset in a neighborhood of  $x$  is used. One important factor for localization is the number of points in that neighborhood. We set  $\alpha \in [0, 1]$  as the percentage of data points in the neighborhood.

<sup>165</sup> For creating the local subset of  $x \in \Omega$  several bounding region can be used:

- $S(x, X, \alpha)$ : the minimum bounding sphere,
- $B(x, X, \alpha)$ : the minimum bounding box,
- $CP(x, X, \alpha)$ : the minimum bounding convex polygon (i.e. convex hull),
- $OB(x, X, \alpha)$ : the minimum oriented bounding box,
- $E(x, X, \alpha)$ : the minimum bounding ellipsoid,

The value of  $\alpha$  for these methods should be considered as a predefined value, and if we set  $\alpha = 1$ , all the data points in  $X$  will lie in the neighborhood.

### 3.3. The Mahalanobis Distance

It is important to use a proper metric for measuring distance in localization. We consider Mahalanobis distance which includes the distribution of the data points. The Mahalanobis distance between a point  $x \in \mathbb{R}^d$  and an arbitrary set of points  $X \subset \mathbb{R}^d$  is defined as [37]

$$D_X(x) = \sqrt{(x - \mu)^T \text{cov}^{-1}(X)(x - \mu)}, \quad (3.7)$$

where  $\text{cov}(X)$  is the covariance matrix of the set  $X$ . The Mahalanobis distance between two points  $y, z \in \mathbb{R}^d$  with respect to the set of points  $X \subset \mathbb{R}^d$  is given as:

$$d_X(y, z) = \sqrt{(y - z)^T \text{cov}^{-1}(X)(y - z)}. \quad (3.8)$$

If there is no specific trend in  $X$  (i.e. no correlation between coordinates of points), then the matrix  $\text{cov}(X)$  becomes a diagonal matrix. Thus, the Mahalanobis distance between points simplified to a scaled Euclidean distance:

$$d(y, z) = \|\Sigma^{-1}(y - z)\|_2, \quad (3.9)$$

where  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_d)$ , in which  $\sigma_i$  is the standard deviation of  $i$ -th variable, for  $i = 1, \dots, d$ .

## <sup>175</sup> 4. Ellipsoidal Quantiles

To approximate a function by a local method at  $x \in \Omega$ , any localization methods described in Section 3.2 can be used to define the neighborhood of  $x$ . For MLS, the neighborhood of  $x$  is a disk  $D(x, \delta)$  (or in general case a hypersphere) with radius  $\delta$  centred at  $x$ . The challenge is that the distribution of points is ignored when we <sup>180</sup> use a hypersphere. In this work, we change that and consider ellipsoidal regions centred at  $x$ . This ellipsoidal neighborhood is constructed based on the trend of local data and it contains points with the most impact on the approximating function (See Figure 4). For constructing the ellipsoidal neighborhoods, we use *quantiles* to include a specific percentage of data [21, 23].

We consider three types of ellipsoidal quantiles. The first one is independent of the location of the point <sup>185</sup> while the others are centred at the given point and constructed from the covariance of data.

### 4.1. Mean-Centric Ellipsoidal Quantiles

Suppose  $\mu$  to be the mean of the points set  $X$ . For the given  $\alpha$  in the interval  $[0, 1]$  we define  $E(X, \alpha)$  as the smallest ellipsoid centred in  $\mu$  where  $\alpha$  is percentile of the points in  $E(X, \alpha)$ .

The directions and magnitudes of radii of the ellipsoid are defined by the covariance of  $X$ . The radii of the ellipsoid are scaled versions of the eigenvalues of  $\text{cov}(X)$ , and the directions of them are defined by the

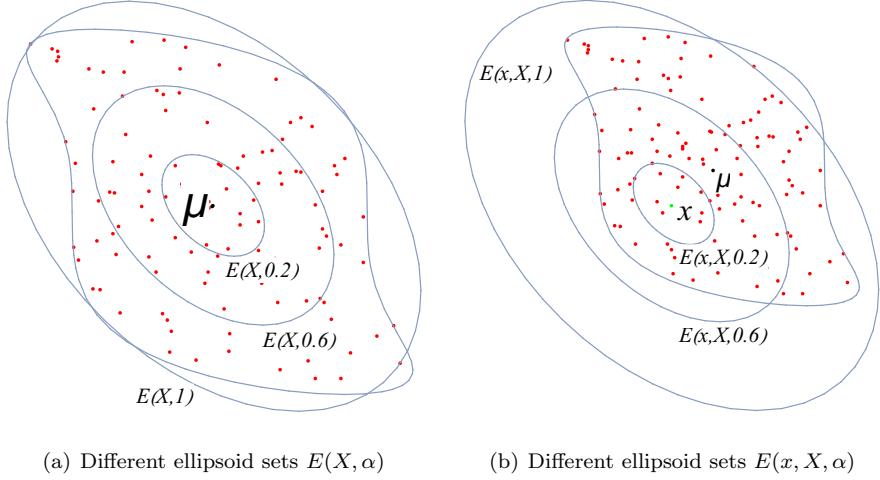


Figure 4: This figure shows the ellipsoidal neighbourhoods of a given set of points for  $\alpha = 0.2, 0.6$  and  $1$ .

eigenvectors of  $\text{cov}(X)$ . If we rank from the largest to smallest eigenvalues, the largest eigenvector shows the direction of the largest radius (i.e. the largest variation in data), the second-largest eigenvector shows the direction of the second radius, and so on. We use Mahalanobis distance  $D_X$  in Equation (3.7) to determine  $E(X, \alpha)$  by

$$E(X, \alpha) = \{y \in \mathbb{R}^d : D_X(y) = C_\alpha\}, \quad (4.1)$$

where  $C_\alpha$  is a constant related to  $\alpha$  that should be determined.

Figure 4(a) shows three different mean-centric ellipsoidal quantiles for  $\alpha = 0.2, 0.6$  and  $1$ .  $E(X, \alpha)$  does not depend on the point  $x$  and the challenge is how to find a point centric ellipsoidal quantile for a given point  $x$ . One may translate  $E(X, \alpha)$  to  $x$  (by  $x - \mu$ ), but this ellipsoid does not capture the distribution of data in the neighborhood of  $x$ .

#### 4.2. Point-Adaptive Ellipsoidal Quantiles

The main goal is to properly approximate the function at  $x \in \Omega$ . Therefore, we introduce a method to centralize the ellipsoidal quantile at  $x$ . For a given  $\alpha$  in the interval  $[0, 1]$ , we define  $E(x, X, \alpha)$  as the smallest ellipsoid centred in  $x$ , where  $\alpha$  is the percentile of the points in this ellipsoidal neighborhood. We use Mahalonobis distance  $d_X$  in Equation (3.8) to determine  $E(x, X, \alpha)$  by

$$E(x, X, \alpha) = \{y \in \mathbb{R}^d : d_X(x, y) = C_\alpha\}, \quad (4.2)$$

where  $C_\alpha$  is a constant related to  $\alpha$ , that should be determined. Figure 4(b) shows three different point-adaptive ellipsoidal quantiles for  $\alpha = 0.2, 0.6$  and  $1$ .

### 4.3. Weighted Point-Adaptive Ellipsoidal Quantiles

To have a proper approximation of the function at  $x \in \Omega$ , it is better to control the impact of the points in the neighborhood of  $x$  (i.e. farther points less and closer points more impact). Therefore, we use a weight function proportional to the inverse of the distance between the data point and  $x$ . We introduce and compare two weighting methods from the mean and covariance. Some types of weighted covariance are applied in Computer Graphics [2, 46]. To define weighted mean and covariance, we apply an auxiliary function  $\psi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  which is a decreasing function. For the given point  $x \in \Omega$ , we extend the definition of mean and covariance by including  $\psi$  as a weight function, capturing the importance of different points in the neighborhood of  $x$ . Also, recall that  $X$  is as an  $n \times d$  matrix, with columns  $X_j = (\gamma_{1,j}, \dots, \gamma_{n,j})^T$ , for  $j = 1, \dots, d$ .

#### 4.3.1. Directional Weighted Covariance

Since the distribution of data points around  $x$  can be directional, the weight function can depend on the axes of the ellipsoid and, consequently, the covariance of  $X$ . For this purpose, the function  $\psi_i$  is considered as a weight function in the  $i$ -th direction. Thus, we define the  $i$ -th component of the generalized weighted mean of  $X$  with respect to  $x$  as follows:

$$\hat{\mu}_i = \frac{\sum_{k=1}^n \gamma_{k,i} \psi_i(|\gamma_i - \gamma_{k,i}|)}{\sum_{k=1}^n \psi_i(|\gamma_i - \gamma_{k,i}|)}, \quad i = 1, \dots, d. \quad (4.3)$$

Also, the weighted covariance of the  $i$  and  $j$  dimensions of  $X$  related to the point  $x$ , can be defined by

$$\text{wcov}_{1,x}(X_i, X_j) = \sum_{k=1}^n \frac{(\gamma_{k,i} - \hat{\mu}_i)(\gamma_{k,j} - \hat{\mu}_j) \sqrt{\psi_i(|\gamma_i - \gamma_{k,i}|)\psi_j(|\gamma_j - \gamma_{k,j}|)}}{\sum_{k=1}^n \sqrt{\psi_i(|\gamma_i - \gamma_{k,i}|)\psi_j(|\gamma_j - \gamma_{k,j}|)} - 1}. \quad (4.4)$$

Therefore, the weighted covariance matrix related to the point  $x$  and with respect to the weight functions is  $\text{wcov}_{1,x}(X) = [\text{wcov}_{1,x}(X_i, X_j)]_{i,j=1,\dots,d}$ .

By defining  $W_i = [\psi_i(|\gamma_i - \gamma_{k,i}|)]_{k=1,\dots,n}$ ,  $V_i = \text{diag}(W_i)$  and  $D_i = W_i^{\frac{1}{2}}$ , Equation (4.4) can be written as

$$\text{wcov}_{1,x}(X)_{i,j} = \frac{(X - \hat{\mu})_i^T D_i D_j (X - \hat{\mu})_j}{\text{trace}(D_i D_j) - 1}, \quad i, j = 1, \dots, d. \quad (4.5)$$

Now, we can define a new Mahalanobis distance between two points  $y$  and  $z$  in  $\mathbb{R}^d$  with respect to the data set  $X$  and the given point  $x$  as follows:

$$d_{X,x}(y, z) = \sqrt{(y - z)^T \text{wcov}_{1,x}^{-1}(X)(y - z)}. \quad (4.6)$$

Furthermore, For a given point  $x \in \Omega$ , we use the new Mahalanobis distance defined in (4.6) to determine a new ellipsoidal quantiles  $GE_1(x, X, \alpha)$  by

$$GE_1(x, X, \alpha) = \{y \in \mathbb{R}^d : \sqrt{(x - y)^T \text{wcov}_{1,x}^{-1}(X)(x - y)} = C_\alpha\}, \quad (4.7)$$

where  $\text{wcov}_{1,x}$  is the weighted covariance matrix defined in (4.4) and  $C_\alpha$  is a constant related to  $\alpha$ , that should be determined.

<sup>215</sup> 4.3.2. Multivariate Weighted Covariance

Same as the previous subsection, in this subsection, we want to specify a weight to each data point such that this weight is proportional to inverse of distance between data point and  $x$ . However, we want to define a single weight function, rather than considering a weight function for each dimension.

By using the function  $\psi$ , we generalize weight functions to multivariate case and we define a weight function  $w : \mathbb{R}^d \rightarrow \mathbb{R}$  as a radial function of  $\psi$ :

$$w(x) = \psi(\|x\|), \quad x \in \mathbb{R}^d. \quad (4.8)$$

In this case, we can define the general mean as a linear combination of the points by using the following equation:

$$\tilde{\mu}_x = \frac{\sum_{i=1}^m x_i w(x - x_i)}{\sum_{i=1}^m w(x - x_i)}. \quad (4.9)$$

This mean value is related to the point  $x$  and based on this mean value, the generalized covariance is defined by:

$$\text{wcov}_{2,x}(X) = \frac{(X - \tilde{\mu}_x)^T W (X - \tilde{\mu}_x)}{\text{trace}(W) - 1}, \quad (4.10)$$

where  $W = \text{diag}(w(x - x_1), \dots, w(x - x_n))$  and  $\text{trace}(W) = \sum_{i=1}^n w(x - x_i)$ .

Same as Equation (4.6), we can define a new Mahalanobis distance between two points  $y$  and  $z$  in  $\mathbb{R}^d$  with respect to the data set  $X$  and the given point  $x$  as follows:

$$d_{X,x}(y, z) = \sqrt{(y - z)^T \text{wcov}_{2,x}^{-1}(X)(y - z)}. \quad (4.11)$$

For a given point  $x \in \Omega$ , we use the new Mahalanobis distance defined in (4.11) to determine a new ellipsoidal quantiles  $GE_2(x, X, \alpha)$  by

$$GE_2(x, X, \alpha) = \{y \in \mathbb{R}^d : \sqrt{(x - y)^T \text{wcov}_{2,x}^{-1}(X)(x - y)} = C_\alpha\}, \quad (4.12)$$

<sup>220</sup> where  $\text{wcov}_{2,x}$  is the weighted covariance matrix defined in (4.10) and  $C_\alpha$  is a constant related to  $\alpha$ , that should be determined.

In summary, for approximating a function at an arbitrary point  $x \in \Omega$  we consider ellipsoids  $E(x, X, \alpha)$ ,  $GE_1(x, X, \alpha)$  and  $GE_2(x, X, \alpha)$  as a local neighborhood of  $x$  in MLS.

**Remark 4.1.** MLS in general, is a deterministic interpolation method, same as IDW, because these methods <sup>225</sup> are based on specific function that determine the smoothness of the resulting surface. The proposed MLS is also lying in this class of interpolation methods. Some methods such as the family of the kriging method are based on statistical models that include autocorrelation. From the other hand, Kriging is a good interpolation technique, if the data is stationary, has a normal distribution and doesn't contain any trend. The advantage of MLS in comparison with Kriging methods is that for MLS, the data does not need to meet the mentioned <sup>230</sup> criteria.

## 5. Model Description on Daily Air Temperature and Humidity

To evaluate and compare the models, we used a core i7-3632 computer with 2.2GHz CPU and 8 GB RAM, and all algorithms were coded in Mathematica 12.1. We considered the region  $\Omega$  to be Alberta Province, 235 Canada. Alberta is a large province with an area of  $661,848\text{km}^2$ . The province extends  $1,223\text{km}$  north to south and  $660\text{km}$  east to west at its maximum width and its climate, varies considerably.

We use a weather dataset for Alberta, Canada. Figure 1 shows the location of weather stations in this province. The data were downloaded from the website of Climate Data Canada (<https://climatedata.ca>). The data observations are limited to the daily average minimum and maximum temperatures and daily average 240 minimum and maximum humidity values.

The data of these stations are used in the period of two years 2018 (January 01 to December 30) and 2019 (January 01 to December 31).

By using a code in Mathematica, the raw data is processed and filtered to extract the coordinates of the stations as well as the maximum and minimum temperature and humidity over the course of days, months and 245 years.

We expand the selected stations by some close out of the province weather stations to approximate the weather function at points near to the border of the province. Therefore, we have collected the data of 196 stations, 120 inside the province, and 76 outside of the province. The minimum and maximum distances between a station and its nearest station are  $1.3\text{km}$  and  $210\text{km}$ , respectively. Also, the furthest distance 250 between two stations is  $1130\text{km}$ . (See Figure 1(c).)

We use a Voronoi-based buffering for the expansion. From the selected region  $\Omega$  (Alberta in our case study), we define the extended region

$$\Omega_{1+\epsilon} = \Omega + \epsilon S_1, \quad (5.1)$$

where  $S_1$  is the unit circle in  $\mathbb{R}^2$ . (In general we consider  $\Omega_{1+\epsilon} = \Omega + \epsilon S_{d-1}$ , where  $S_{d-1}$  is the unit sphere surface in  $\mathbb{R}^d$ .) Now, let  $X^G$  be the set of all data points (stations) and  $X$  be the set of all nodal points in the region  $\Omega$ . i.e.  $X = X^G \cap \Omega$ . So the set of all data points (stations) in  $\Omega_{1+\epsilon}$  is  $X_{1+\epsilon}$ . For every  $x \in \Omega$ , the ellipsoids can be built over  $X_{1+\epsilon}$ :

$$GE_i(x, X_{1+\epsilon}, \alpha) = \{y \in \mathbb{R}^2 : d_{X_{1+\epsilon}, x}(x, y) = 1\}, \quad x \in \Omega, \quad i = 1, 2. \quad (5.2)$$

In this case, there is more information on the points near the boundary. Figure 5, shows  $\Omega_{1+\epsilon}$  and a new ellipsoid for a specific point  $x \in \Omega$ .

In order to specify a value for  $\epsilon$ , we can consider the smallest positive number  $\epsilon$ , such that the Voronoi diagram of all points in  $\Omega_{1+\epsilon}$  covers  $\Omega$ . i.e.  $\Omega$  be a subset of Voronoi diagram of  $\Omega_{1+\epsilon}$ . For the weather stations in

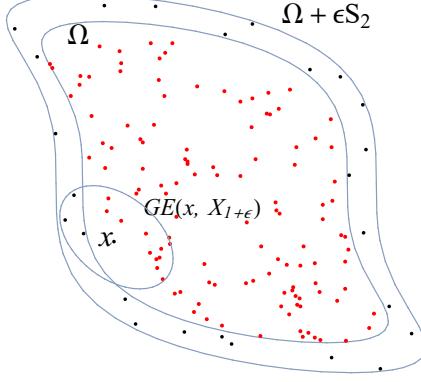


Figure 5: A set of points and a band around it and the new ellipsoidal for the point  $x$ .

<sup>255</sup> Alberta see Figure 6.

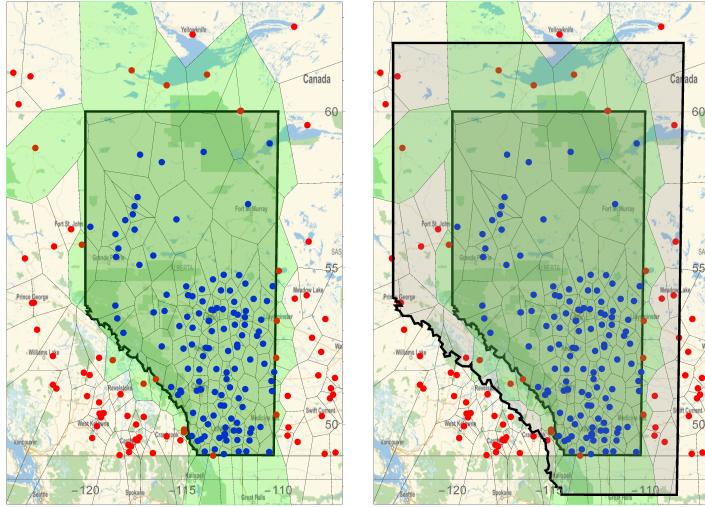
Based on Remark 4.1, it is better to compare the proposed method with IDW. Also, since KNN is a widely used method with lots of applications, we combined KNN with both MLS and IDW in this section. The MLS weight function  $\phi$  is considered positive inside the unit interval  $[0, 1]$  and zero elsewhere. Also, the degree of bivariate polynomials is chose to be 2, so the polynomial basis is  $\{1, \gamma_1, \gamma_2, \gamma_1^2, \gamma_1\gamma_2, \gamma_2^2\}$ .

<sup>260</sup> In this comparison, we used local MLS over 6 neighborhood sets:

- $E(x, X, \alpha)$  using (4.2), with a specific  $\alpha$  such that  $C_\alpha = 1$ ,
- Insphere or interior disk of  $E(x, X, \alpha)$ ,
- Circumsphere or Outer disk of  $E(x, X, \alpha)$ ,
- $GE_1(x, X, \alpha)$  using (4.5) and (4.7), with a specific  $\alpha$  such that  $C_\alpha = 1$ ,
- <sup>265</sup>  $GE_2(x, X, \alpha)$  using (4.10) and(4.12), with a specific  $\alpha$  such that  $C_\alpha = 1$ ,
- a region containing  $k$  nearest points.

and we named them MLS-E, MLS-I, MLS-O, MLS- $GE_1$ , MLS- $GE_2$  and MLS-KNN, respectively. Figure 7 shows the sets E, I and O for an arbitrary point  $x \in \Omega$ . Another method which used in this comparison is IDW. We applied both global IDW which uses all data points and a local IDW with  $k$  nearest points, named IDW-KNN.

<sup>270</sup> In the methods MLS-E, MLS-I, MLS-O, MLS- $GE_1$  and MLS- $GE_2$ , the number of points located in the ellipsoids or disks, varies for different stations, so we used the average number of points used in them as a fixed number of neighborhood points in computations for KNN methods. i.e. In two methods MLS-KNN and IDW-KNN,  $k$  is the average number of points used in MLS-E, MLS-I, MLS-O, MLS- $GE_1$  and MLS- $GE_2$ . The



(a) Voronoi diagram of stations which are needed for all Alberta province weather computation.  
(b) Alberta and a band around it as which are needed for all Alberta  $\Omega_{1+\epsilon}$ .

Figure 6: Detecting the useful points for computation.

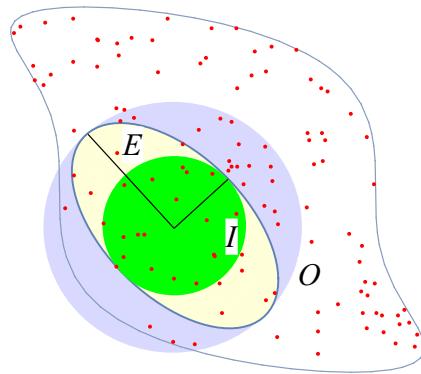


Figure 7: Different neighborhoods of a points: An ellipse  $E$ , its insphere  $I$  and its outer disk  $O$ .

results obtained from  $k$  nearest neighbors with both IDW and MLS interpolation methods are used to have a  
275 better comparison between the methods.

Algorithm 1 describes the steps of ellipsoidal MLS computation.

---

**Algorithm 1** Ellipsoidal MLS

---

**Input:** A set of data points  $\{x_i, f_i\}_{i=1}^n \in \mathbb{R}^{d+1}$ , and  $x \in \Omega$ .

**Output:** The approximating value for  $f(x)$ .

- 280 1. Compute the mean of data (traditional mean, weighted means form (4.3), or (4.9)).  
2. Compute the covariance of data (traditional covariance, weighted covariance form (4.4), or (4.10)).  
3. Determine the ellipsoid related to the computed covariance. For MLS-I and MLS-O, we consider disks  
with the minimum and maximum radius of the ellipsoid, respectively.  
4. Compute the points lie in the related ellipsoid.  
285 5. Apply MLS on the obtained points using (3.5) and (3.6).
- 

The main difference between the proposed method and the traditional MLS method is consideration of the ellipsoidal neighborhood instead of disk or ball one. Also, in this method, the radii of the ellipsoid are the deviations of data, rather than a predefined positive number. The only difference between the computational cost of this method and traditional MLS is the ways of choosing local points. The extra computation we need  
290 for this method with respect to MLS is the computation of the mean (a 2D vector) and the inverse of the covariance matrix (2 by 2). Therefore, approximating by these methods has  $O(n)$  more computational cost than traditional MLS.

## 6. The Results and Discussion

295 To assess the performance of each method, we exclude one of the stations and approximate the function values (temperature and humidity) at the location of that station from the approximating function. The value of the function at a station is approximated by using the values of the function on the other stations. The aggregation of approximation errors over all points were expressed by RMS, MAE, RMSP, and MAEP as four criteria.

300 In this section, we present the results of applying MLS in different types of neighborhoods. To have the same situation, the Gaussian function used for all MLS methods with a parameter 0.01.

To compare the results, we used two quantities; Root Mean Squares (RMS) Error and Mean of Absolute Errors (MAE). These are defined as:

$$\text{RMS} = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}}, \quad (6.1)$$

and

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n e_i, \quad (6.2)$$

where  $n$  is the number of all points and  $e_i$  is the absolute error for  $i$ th point.

We also apply a quantity named “Pointwise Error Effect” to compare the methods based on both their neighborhood points and the obtained errors:

$$pe_i = p_i \cdot e_i, \quad (6.3)$$

in which  $e_i$  is the error of the method on  $i$ th point and  $p_i$  is the percentage of the number of points in the local domain in order to obtain the results. So, for two error quantities RMS and MAE, we also computed RMSP and MAEP, which are RMS and MAE of  $pe_i$ , respectively, and show the performance of error with respect to the percentile of all used points in the computation.

---

Algorithm 2 shows the steps of evaluation of the methods.

**Algorithm 2** Evaluation by Ellipsoidal MLS

---

**Input:** A set of data points  $\{x_i, f_i\}_{i=1}^n \in \mathbb{R}^{d+1}$ , and  $x \in \Omega$ .

**Output:** The error of the method.

1. Considering a fix index  $i$ , do the following:
    - (a) Omit the  $i$ th point:  $X_i = \{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$
    - (b) Approximate the value of function over  $X_i$  and name it  $s_i$ .
    - (c) Compute the error:  $e_i = f_i - s_i$ .
  2. Compute RMS, MAE, RMSP and MAEP.
- 

In Tables 1 and 2, RMS and MAE of the methods for the approximated values of temperature of the whole years 2018 and 2019 are shown. In Tables 3 and 4, also RMSP and MAEP of the methods for the same periods are given.

In Figures 8 and 9, the diagrams of RMS and MAE of the methods for the approximated values of temperature of the whole years 2018 and 2019 are shown. In Figures 10 and 11, also RMSP and MAEP of the methods for the same period are given.

Figures 12 and 13, shows RMS and MAE of the local methods for the approximated values of temperature of the whole years 2018 and 2019 period. Figures 14 and 15, also shows RMSP and MAEP of these methods for the same periods.

In Table 5, RMS and MAE of the methods and in Table 6, RMSP and MAEP of the methods for the approximated values of humidity of the whole year 2019 are given.

Also, in Figure 16, RMS and MAE of the methods and in Figure 17, RMSP, and MAEP of the methods for the approximated values of humidity of the whole year 2019 are shown.

	RMS-T <sub>min</sub>	RMS-T <sub>max</sub>	MAE-T <sub>min</sub>	MAE-T <sub>max</sub>
MLS-E	2.089	1.674	1.590	1.175
MLS-GE <sub>1</sub>	2.128	1.654	1.606	1.166
MLS-GE <sub>2</sub>	2.079	1.640	1.575	1.157
MLS-I	3.457	2.615	1.867	1.334
MLS-O	2.306	1.913	1.645	1.237
MLS-KNN	2.084	1.755	1.580	1.167
IDW-KNN	2.260	1.984	1.712	1.342
IDW	2.398	2.185	1.794	1.482

Table 1: MAE and RMS for T<sub>min</sub> and T<sub>max</sub> of year 2018.

	RMS-T <sub>min</sub>	RMS-T <sub>max</sub>	MAE-T <sub>min</sub>	MAE-T <sub>max</sub>
MLS-E	1.839	1.566	1.312	1.039
MLS-GE <sub>1</sub>	1.923	1.611	1.350	1.051
MLS-GE <sub>2</sub>	1.882	1.636	1.330	1.058
MLS-I	2.760	2.068	1.526	1.158
MLS-O	1.902	1.803	1.317	1.085
MLS-KNN	1.784	1.673	1.279	1.050
IDW-KNN	1.878	1.830	1.356	1.206
IDW	2.008	2.096	1.442	1.416

Table 2: MAE and RMS for T<sub>min</sub> and T<sub>max</sub> of year 2019.

	RMSP-T <sub>min</sub>	RMSP-T <sub>max</sub>	MAEP-T <sub>min</sub>	MAEP-T <sub>max</sub>
MLS-E	1.077	0.865	0.820	0.607
MLS-GE <sub>1</sub>	0.884	0.688	0.668	0.486
MLS-GE <sub>2</sub>	0.956	0.755	0.725	0.533
MLS-I	1.596	1.209	0.861	0.616
MLS-O	1.323	1.098	0.943	0.710
MLS-KNN	1.013	0.855	0.769	0.569
IDW-KNN	1.097	0.965	0.832	0.653
IDW	2.398	2.185	1.794	1.482

Table 3: MAEP and RMSP for T<sub>min</sub> and T<sub>max</sub> of year 2018.

	RMSP-T <sub>min</sub>	RMSP-T <sub>max</sub>	MAEP-T <sub>min</sub>	MAEP-T <sub>max</sub>
MLS-E	0.875	0.744	0.624	0.494
MLS-GE <sub>1</sub>	0.726	0.607	0.510	0.396
MLS-GE <sub>2</sub>	0.784	0.680	0.554	0.440
MLS-I	1.146	0.857	0.634	0.481
MLS-O	1.015	0.961	0.703	0.578
MLS-KNN	0.789	0.739	0.566	0.464
IDW-KNN	0.831	0.809	0.601	0.534
IDW	2.008	2.096	1.442	1.416

Table 4: MAEP and RMSP for T<sub>min</sub> and T<sub>max</sub> of year 2019.

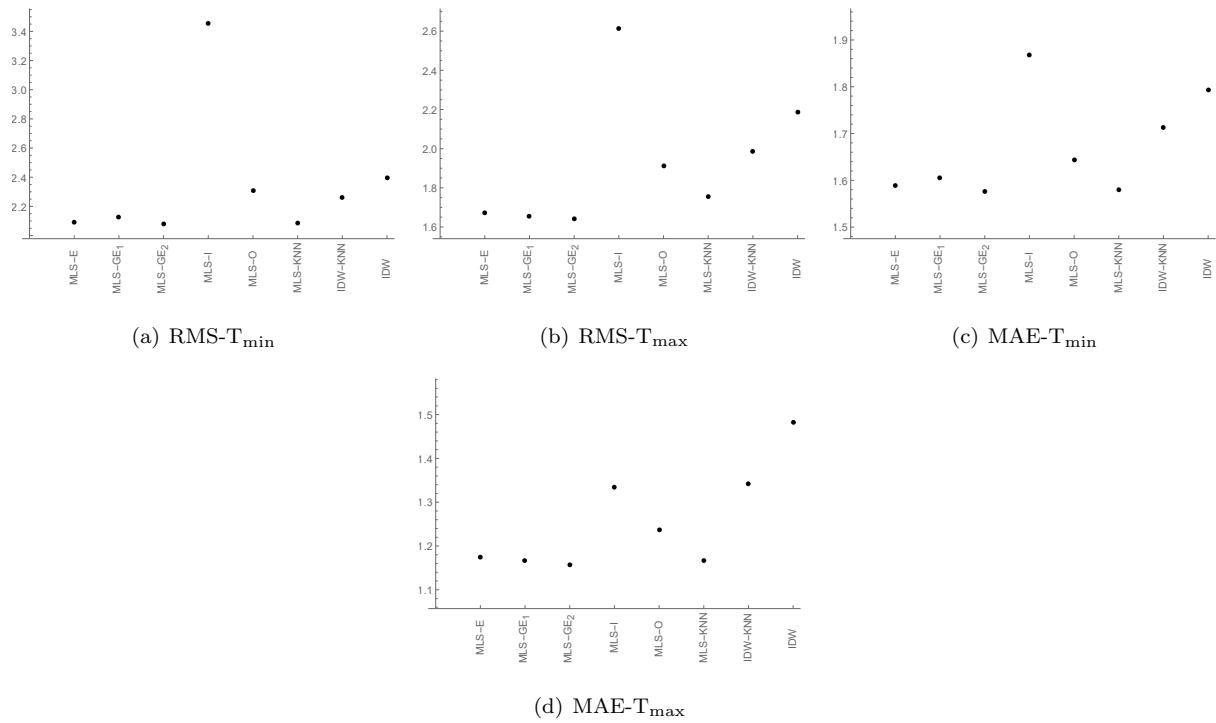


Figure 8: MAE and RMS for T<sub>min</sub> and T<sub>max</sub> for year 2018.

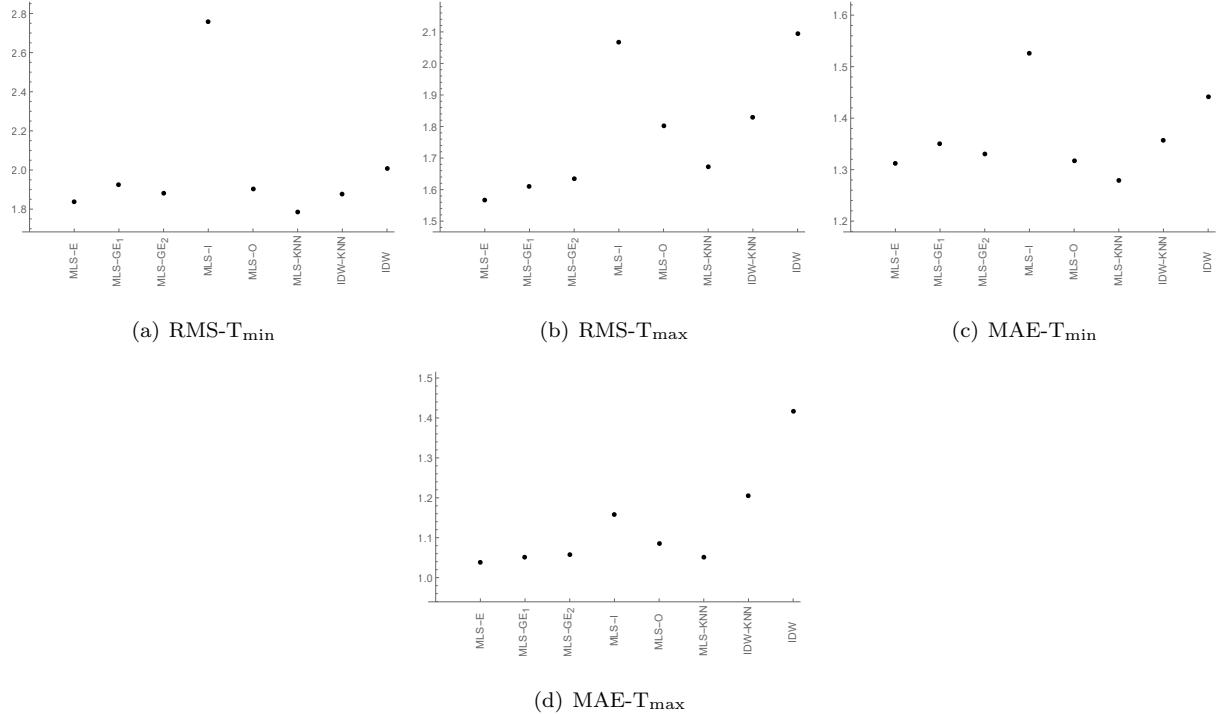


Figure 9: MAE and RMS for  $T_{\min}$  and  $T_{\max}$  for year 2019.

	RMS-H <sub>min</sub>	RMS-H <sub>max</sub>	MAE-H <sub>min</sub>	MAE-H <sub>max</sub>
MLS-E	3.706	2.718	2.766	1.988
MLS-GE <sub>1</sub>	3.838	2.801	2.778	2.035
MLS-GE <sub>2</sub>	3.730	2.816	2.755	2.033
MLS-I	4.464	3.298	2.970	2.165
MLS-O	3.893	2.776	2.819	1.995
MLS-KNN	3.739	2.733	2.756	1.957
IDW-KNN	4.010	2.794	3.012	2.004
IDW	4.299	2.892	3.261	2.058

Table 5: MAE and RMS for  $H_{\min}$  and  $H_{\max}$  of year 2019.

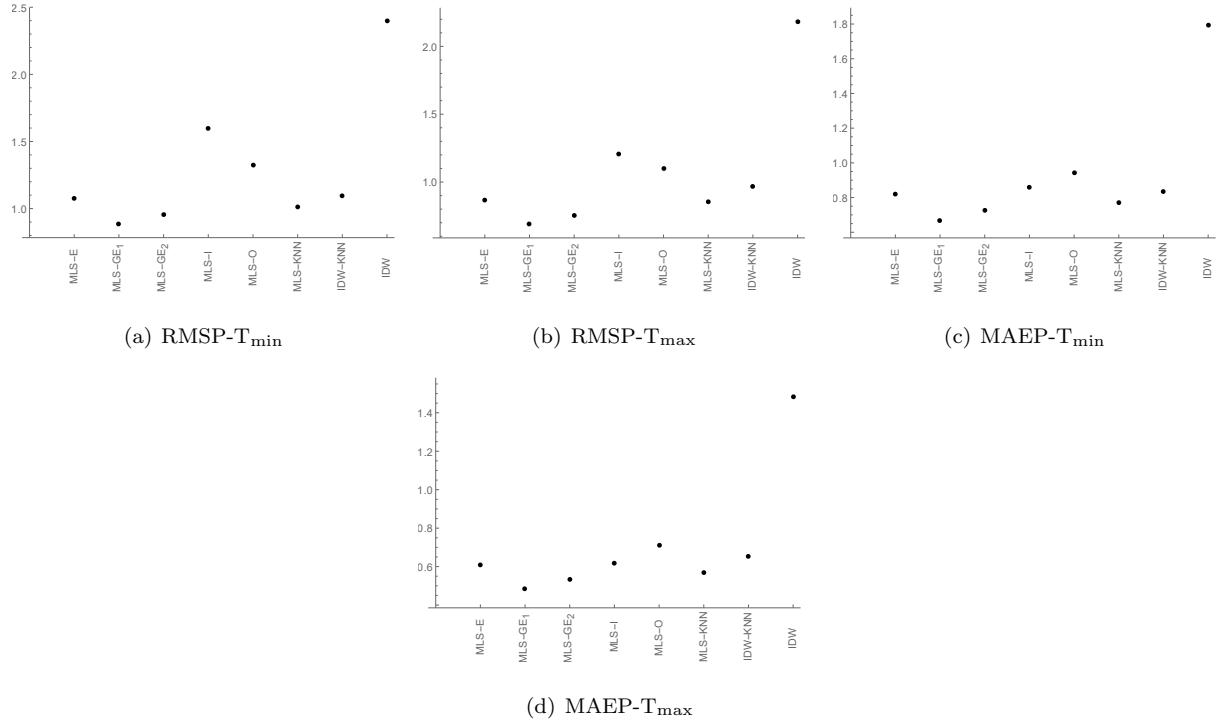


Figure 10: MAEP and RMSP for T<sub>min</sub> and T<sub>max</sub> for year 2018.

	RMSP-H <sub>min</sub>	RMSP-H <sub>max</sub>	MAEP-H <sub>min</sub>	MAEP-H <sub>max</sub>
MLS-E	1.764	1.293	1.317	0.945
MLS-GE <sub>1</sub>	1.447	1.057	1.048	0.767
MLS-GE <sub>2</sub>	1.554	1.173	1.149	0.846
MLS-I	1.852	1.368	1.233	0.898
MLS-O	2.073	1.481	1.502	1.064
MLS-KNN	1.652	1.209	1.219	0.865
IDW-KNN	1.773	1.236	1.333	0.886
IDW	4.299	2.892	3.261	2.058

Table 6: MAEP and RMSP for H<sub>min</sub> and H<sub>max</sub> of year 2019.

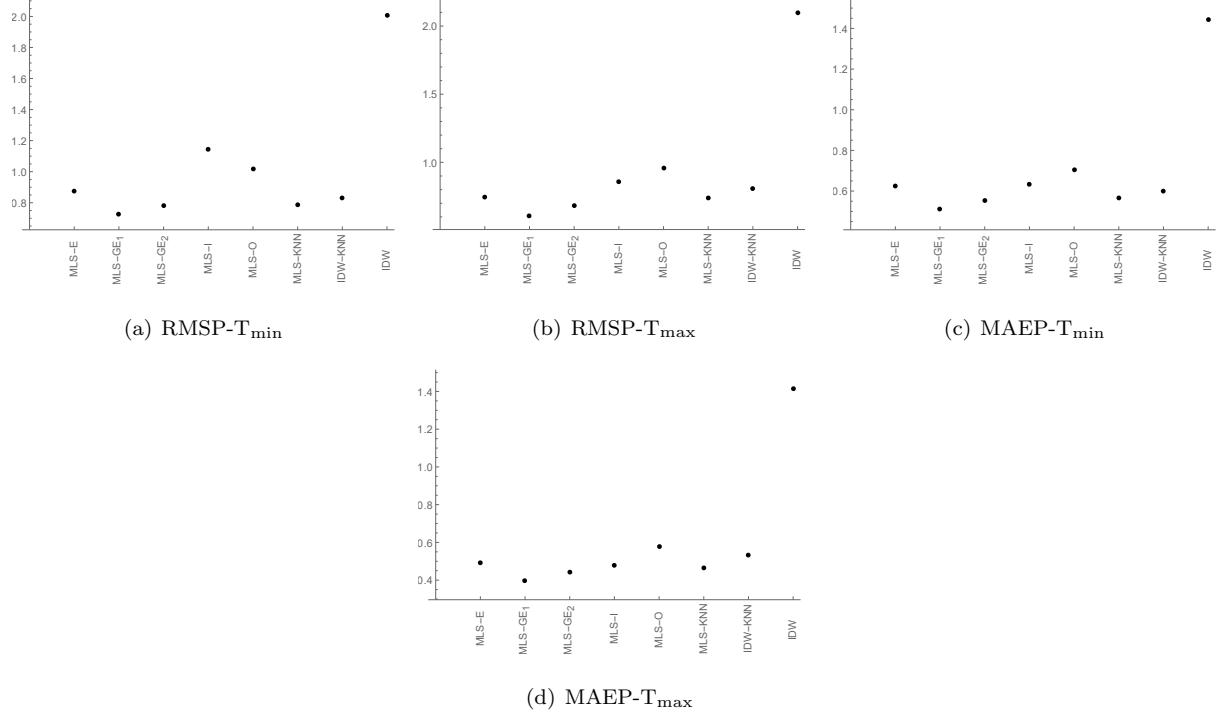


Figure 11: MAEP and RMSP for  $T_{\min}$  and  $T_{\max}$  for year 2019.

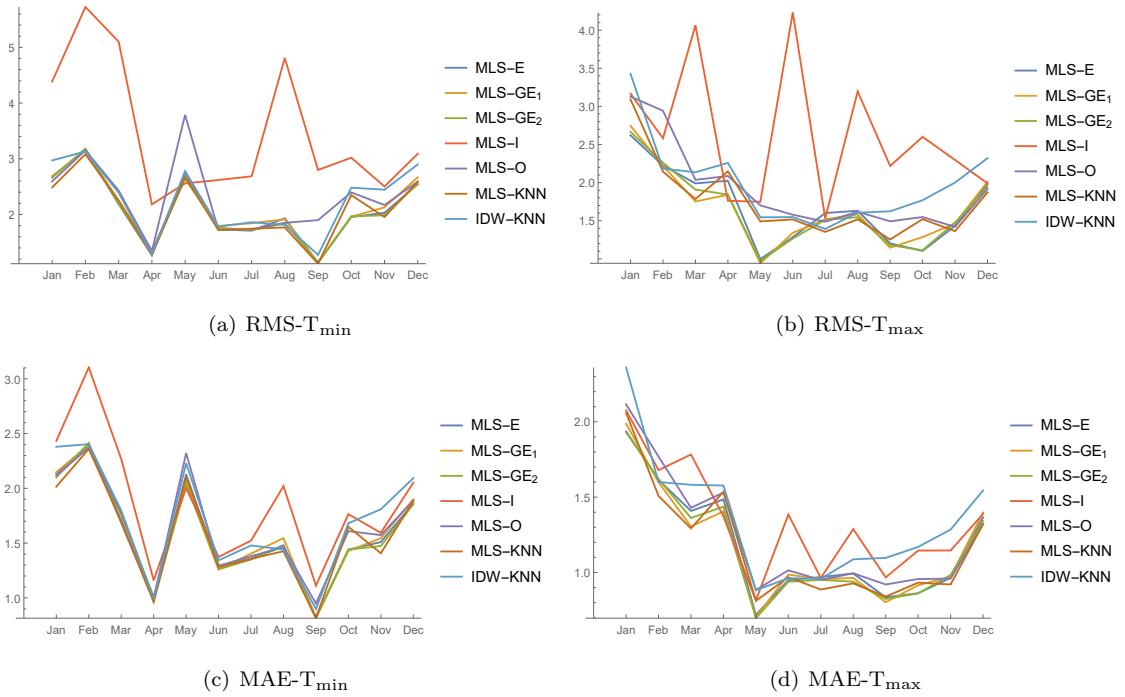


Figure 12: MAE and RMS for  $T_{\min}$  and  $T_{\max}$  for year 2018.

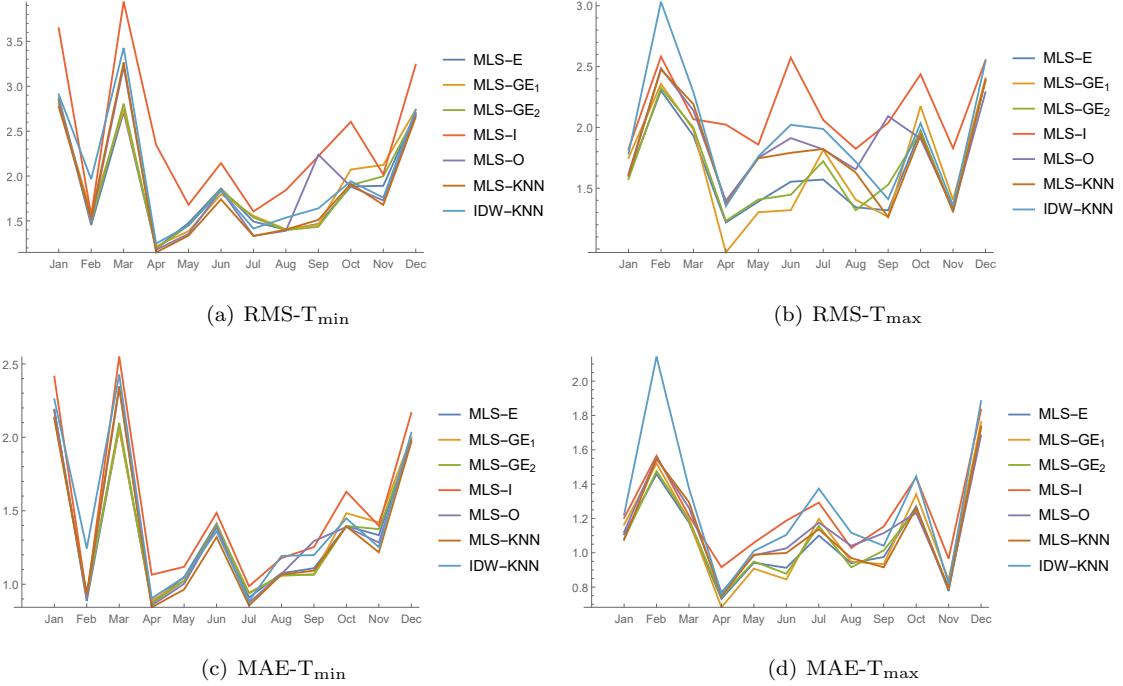


Figure 13: MAE and RMS for  $T_{\min}$  and  $T_{\max}$  for year 2019.

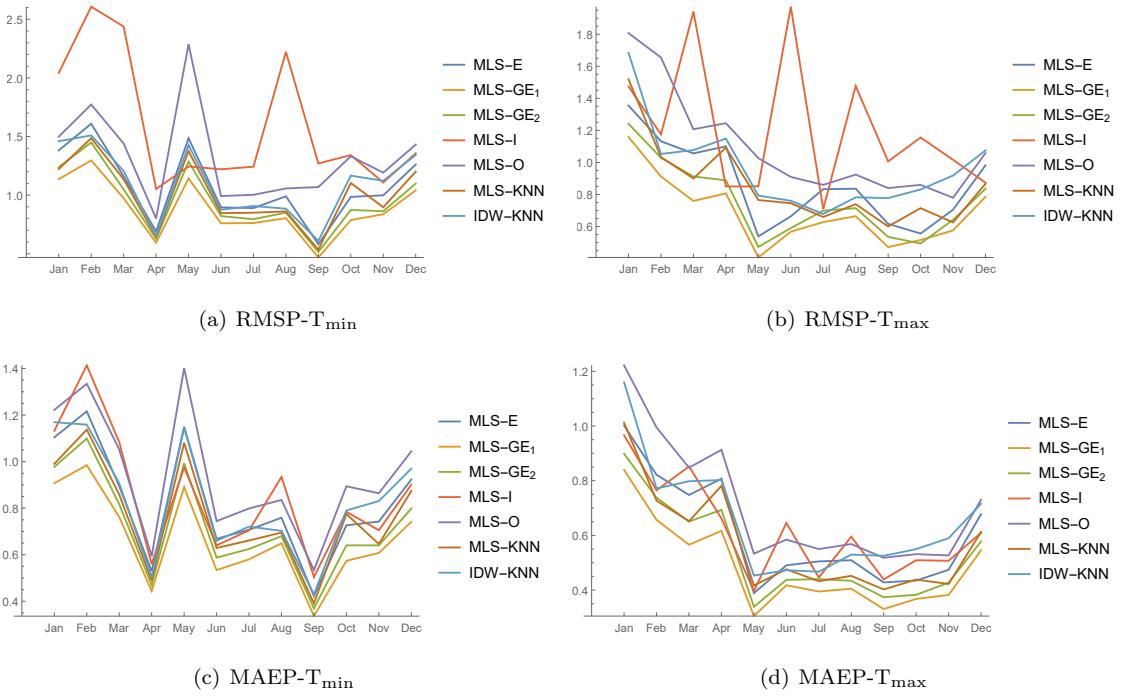


Figure 14: MAEP and RMSP for  $T_{\min}$  and  $T_{\max}$  for year 2018.

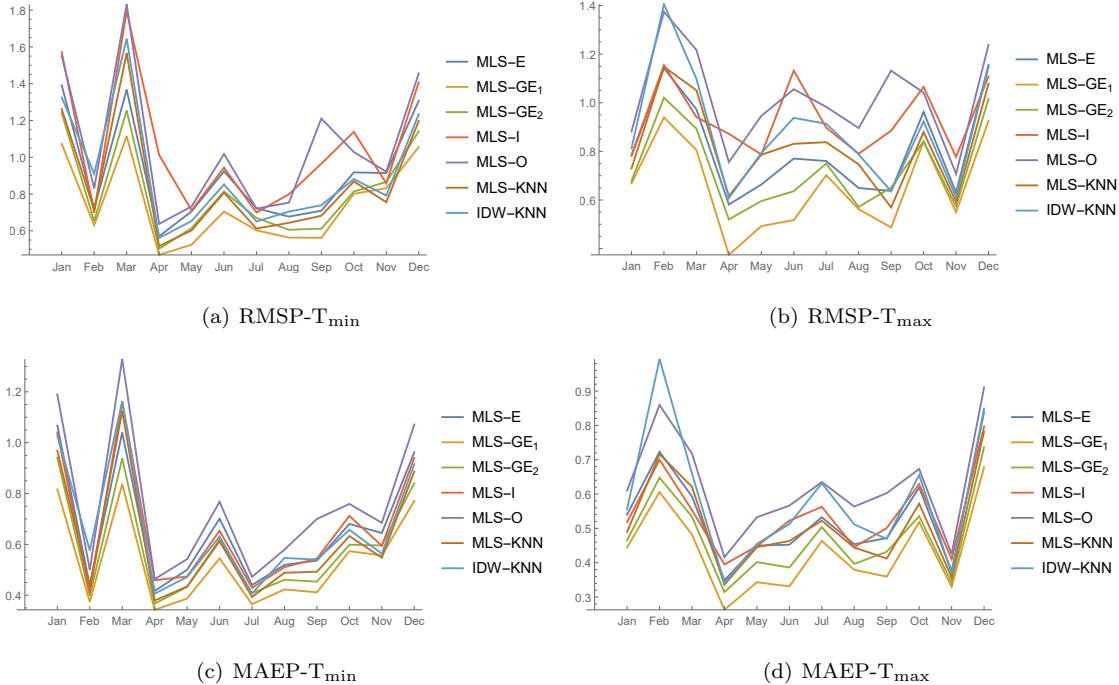


Figure 15: MAEP and RMSP for  $T_{\min}$  and  $T_{\max}$  for year 2019.

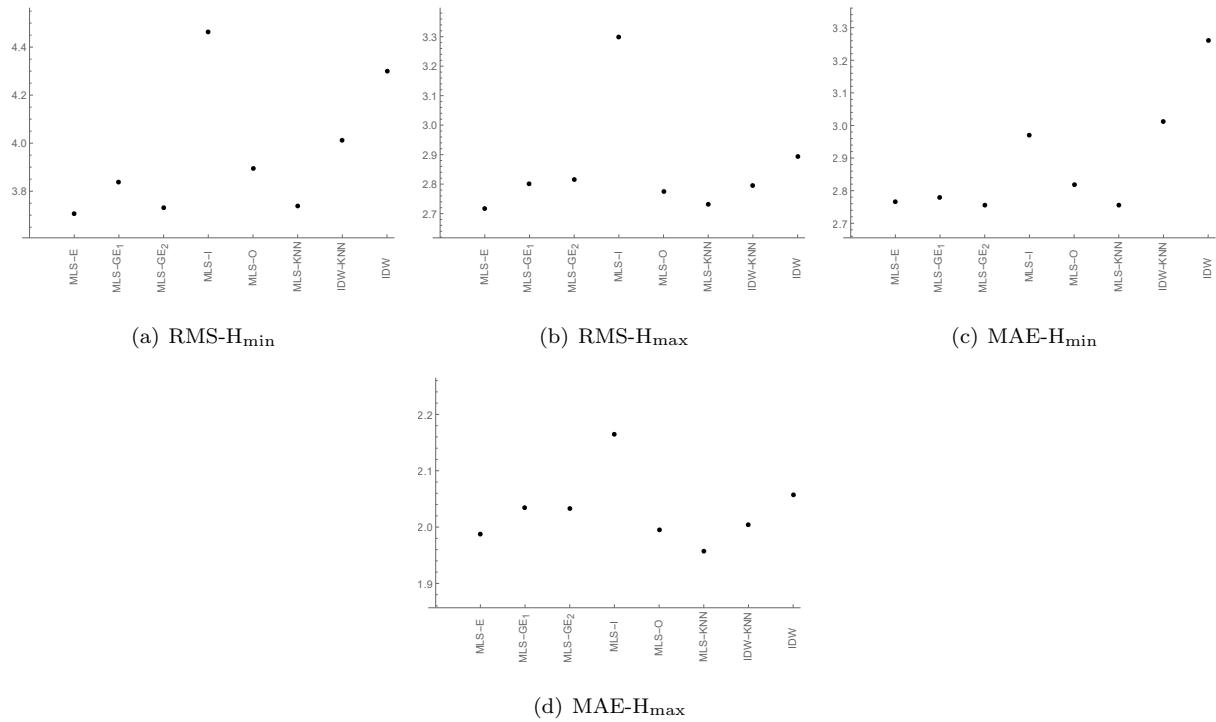


Figure 16: MAE and RMS for  $H_{\min}$  and  $H_{\max}$  for year 2019.

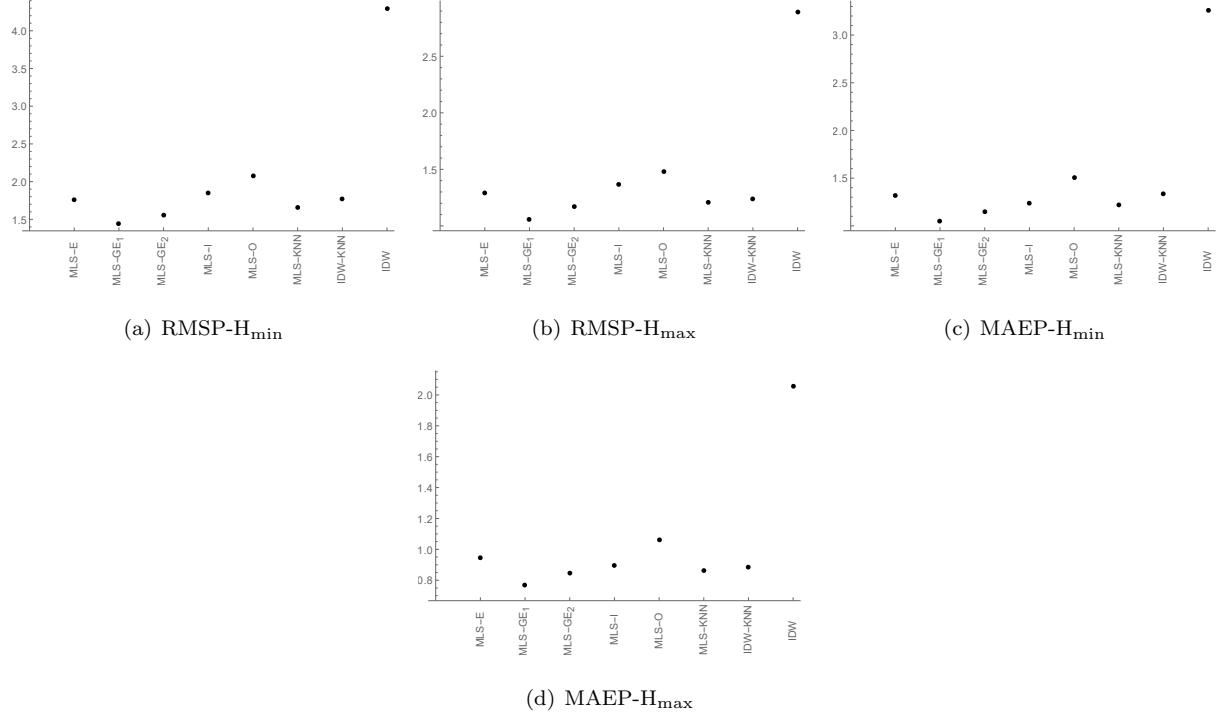


Figure 17: MAEP and RMSP for  $H_{\min}$  and  $H_{\max}$  for year 2019.

Figure 18 shows the heat-map of the results for Max-Temperature on Feb 01, 2019. In this figure blue color shows the smaller error and red color shows the larger error.

Based on the results of Tables 3, 4 and 6, among all methods, MLS-GE<sub>1</sub> has the best performances in  $T_{\min}$ ,  $T_{\max}$ ,  $H_{\min}$  and  $H_{\max}$ , for both RMSP and MAEP. The method MLS-GE<sub>2</sub> has also very good performances in  $T_{\min}$ ,  $T_{\max}$ ,  $H_{\min}$  and  $H_{\max}$ , for both RMSP and MAEP. MLS-KNN and MLS-E, have also better performance than IDW based on RMS and MAE results.

Moreover, from the results, it is clear that the local methods all have better performance than the global method IDW except MLS-I. This is expected because of Remark 3.1, as some of the inspheres did not contain enough data points. For the situation that the moment matrix is singular, we use the least-squares method to solve the given system of equations.

The comparison between the run-time of these methods for finding the stations in the neighborhood of a given point for all Alberta weather stations is given in Table 7

We, also computed the value of a function(minimum temperature) at all the Alberta stations for a day, using the given methods. The computation time for different methods is given in Table 8.

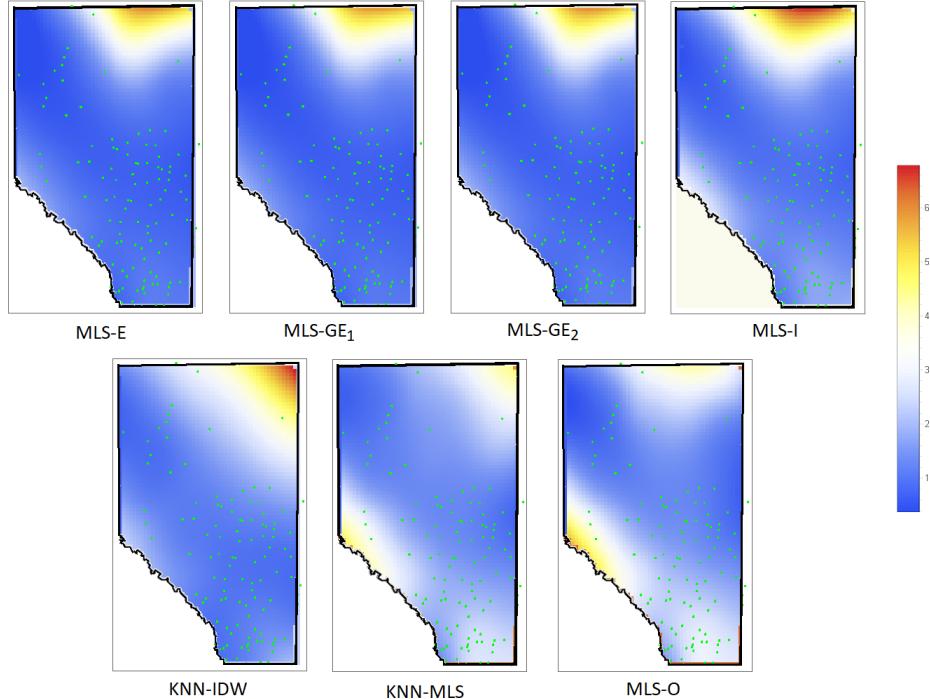


Figure 18: Heat Map of Absolute Error of Max-Temperature in Feb 01, 2019.

MLS-I	MLS-O	MLS-E	MLS-GE <sub>1</sub>	MLS-GE <sub>1</sub>	KNN-MLS	KNN-IDW	IDW
0.080071	0.102345	0.388112	2.111604	0.997865	0.096014	0.097517	—

Table 7: The computation time for locating all neighborhood stations of all Alberta stations (Second).

MLS-I	MLS-O	MLS-E	MLS-GE <sub>1</sub>	MLS-GE <sub>1</sub>	KNN-MLS	KNN-IDW	IDW
3.424717	4.606448	3.838209	4.035241	4.083312	4.084423	4.060301	18.471395

Table 8: The computation time for evaluation of a function over all Alberta stations (Second).

## 7. Conclusions

In this paper, we used the Moving Least Squares method with two new local ellipsoid domains, rather than the disk domain, by using a new method based on weighted covariance and Mahalanobis distance between points. We applied these methods over the problem of daily temperature and humidity approximation over Alberta, Canada. Based on the results, these types of neighborhoods are useful, and MLS has acceptable behavior over them.

## Funding

This research was partially funded by Central Tehran Branch, Islamic Azad University, and it was partially  
355 funded by Natural Sciences and Engineering Research Council of Canada (NSERC).

## Acknowledgement

The authors wish to thank the anonymous referees for their valuable comments and their useful suggestion which improved the quality of this paper. We also, wish to thank Parsa Samavati and Benjamin Ulmer for checking English grammar and punctuation of this paper.

## 360 Conflicts of Interest

The authors declare no conflict of interest.

## References

- [1] Adams, B., Wicke M., (2009), Meshless Approximation Methods and Applications in Physics Based Modeling and Animation, Editors: K. Museth and D. Weiskopf, 2009, The Eurographics Association. DOI: 10.2312/egt.20091070
- 365 [2] Adamson, A., Alexa, M., Approximating and intersecting surfaces from points. In: SGP '03: Proceedings of the 2003 Eurographics, ACM SIGGRAPH symposium on Geometry processing, 230-239. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland (2003).
- [3] Ahrens B., Distance in spatial interpolation of daily rain gauge data, *Hydrol. Earth Syst. Sci.*, 10 (2006) 197–208.
- 370 [4] Altman N., An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *The American Statistician*, 46(3) (1992) 175-185. DOI:10.2307/2685209.
- [5] Amirkhalian M., Arghand M., Kansa E.J., A new approximate method for an inverse time-dependent heat source problem using fundamental solutions and RBFs, *Engineering Analysis with Boundary Elements*, 64 (2016) 278-289. DOI:10.1016/j.enganabound.2015.12.016.
- 375 [6] Amirkhalian M., Khademi A., Neisy A., A Collocation Method by Moving Least Squares Applicable to European Option Pricing, *Journal of Interpolation and Approximation in Scientific Computing* 2016 No.1 (2016) 58-65. DOI:10.5899/2016/jiasc-00105.
- [7] Amirkhalian M., Mafkandi H., Approximation of parametric curves by Moving Least Squares method, *Applied Mathematics and Computation*, 283 (2016), 290-298. DOI:10.1016/j.amc.2016.02.039.

- [8] Angulo-Martínez M., López-Vicente M., Vicente-Serrano S. M., Beguería, S., Mapping rainfall erosivity at a regional  
380 scale: a comparison of interpolation methods in the Ebro Basin (NE Spain), *Hydrology and Earth System Sciences*, 13 (10) (2009) 1907-1920. DOI:10.5194/hess-13-1907-2009.
- [9] Banik Sh., Anwer M., Khodadad Khan A.F.M., Soft Computing Models to Predict Daily Temperature of Dhaka,  
Proceedings of 13th International Conference on Computer and Information Technology (ICCIT 2010), (2010),  
Dhaka, Bangladesh.
- 385 [10] Bartels, R. H. and Golub, G. H. and Samavati, F. F., Some observations on local least squares, *BIT Numerical Mathematics*, 46 (3) (2006) 455-477. DOI:10.1007/s10543-006-0075-y.
- [11] Bartels R. H., Samavati F. F., Reversing subdivision rules: local linear conditions and observations on inner products, *Journal of Computational and Applied Mathematics*, 119 (2000) 29-67. DOI:10.1016/S0377-0427(00)00370-8.
- 390 [12] Bhattacharjee Sh., Ghosh S. K., Spatial Interpolation to Predict Missing Attributes in GIS Using Semantic Kriging, *IEEE Transactions on Geoscience and Remote Sensing*, 52 (8) (2014) 4771-1780.
- [13] Bjerknes V., Das Problem der Wettervorhersage betrachtet vom Standpunkt der Mechanik und Physik. *Meteorol. Z.*, 21 (1904) 1-7.
- 395 [14] Borges P. d. A., Franke J., da Anunciação Y. M. T. et al. Comparison of spatial interpolation methods for the estimation of precipitation distribution in Distrito Federal, Brazil. *Theoretical and Applied Climatology*, 123 (2016) 335–348. DOI:10.1007/s00704-014-1359-9.
- [15] Cavoretto, R., Schneider, T., Zulian, P. OPENCL Based Parallel Algorithm for RBF-PUM Interpolation. *Journal of Scientific Computing*, 74 (2018) 267–289. DOI:10.1007/s10915-017-0431-x.
- [16] Chen F.-W., Liu C.-W., Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan, *Paddy and Water Environment*, 10(3) (2012) 209-222. DOI: 10.1007/s10333-012-0319-1.
- 400 [17] Cheney E. W., Light W. A. (2000). *A course in approximation theory*. Pacific Grove: Brooks/Cole Pub. Co.
- [18] Cover T. M., Hart P. E., Nearest neighbor pattern classification, *IEEE Transaction of Information Theory*, 13 (1967) 21-27.
- [19] Dalmau R., Perez-Batlle M., Prats X., Estimation and prediction of weather variables from surveillance data using spatio-temporal Kriging, Conference: 2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC). DOI:  
405 10.1109/DASC.2017.8102132.
- [20] Fasshauer G., McCourt M., Kernel-based Approximation Methods using MATLAB, WORLD SCIENTIFIC, (2015). DOI:10.1142/9335.
- [21] Furno M., Vistocco D., (2018) Quantile Regression: Estimation and Simulation, Wiley Series in Probability and Statistics.

- 410 [22] George Y. Lu, David W. Wong, An adaptive inverse-distance weighting spatial interpolation technique, Computers & Geosciences, 34 (9) (2008), 1044-1055. DOI:10.1016/j.cageo.2007.07.010.
- [23] Hlubinka D., Šiman M., On elliptical quantiles in the quantile regression setup, Journal of Multivariate Analysis, 116 (2013) 163-171. DOI:10.1016/j.jmva.2012.11.016.
- 415 [24] Hofstra N., Haylock M., New M., Jones, P., Frei C., Comparison of six methods for the interpolation of daily, European climate data, Journal of Geophysical Research: Atmospheres 113 (2008) JD010100. DOI:10.1029/2008JD010100.
- [25] Hristopulos D. T., Agou V. D., Stochastic local interaction model with sparse precision matrix for space–time interpolation, Spatial Statistics, (2019) 100403, in press. DOI:10.1016/j.spasta.2019.100403.
- 420 [26] Jain P., Flannigan M., Comparison of methods for spatial interpolation of fire weather in Alberta, Canada, Canadian Journal of Forest Research 47 (12) (2017) 1646-1658. DOI: 10.1139/cjfr-2017-0101.
- [27] Kalnay, E. (2002). Atmospheric Modeling, Data Assimilation and Predictability. Cambridge: Cambridge University Press. DOI:10.1017/CBO9780511802270.
- 425 [28] Kleiber W., Katz R. W., and Rajagopalan B., Daily Minimum and Maximum Temperature Simulation Over Complex Terrain, The Annals of Applied Statistics, 7 (1) (2013) 588-612.
- [29] Koh T.-Y., Ng J. S., Improved diagnostics for NWP verification in the tropics, J. Geophys. Res., 114 (2009) D12102, DOI:10.1029/2008JD011179.
- 430 [30] Koh T.-Y., Wang S., Bhatt B. C., A diagnostic suite to assess NWP performance, J. Geophys. Res., 117 (2012) D13109, DOI:10.1029/2011JD017103.
- [31] Lancaster P., Salkauskas K., Surfaces generated by moving least squares methods, Mathematics of Computation, 37 (1981) 141–158.
- 435 [32] Lazzaro D., Montefusco L. B., Radial basis functions for the multivariate interpolation of large scattered data sets, Journal of Computational and Applied Mathematics 140 (2002) 521–536. DOI:10.1016/S0377-0427(01)00485-X.
- [33] Levin D., Mesh-Independent Surface Interpolation. In: Brunnett G., Hamann B., Müller H., Linsen L. (eds) Geometric Modeling for Scientific Visualization. Mathematics and Visualization. Springer, Berlin, Heidelberg, (2004) 37-49. DOI: 10.1007/978-3-662-07443-5\_3.
- 440 [34] Li S., Yao L., Yi S., Wang W., A Meshless Radial Basis Function Based on Partition of Unity Method for Piezoelectric Structures, Mathematical Problems in Engineering, Volume 2016, Article ID 7632176, 17 pages. DOI:10.1155/2016/7632176.

- [35] Lin J.-F., Liaw Y.-C., Chan K.-Y., Wu S.-C., Fast exact local-mean based classification algorithm using class rejection and initial distance assigning processes, *Pattern Recognition Letters*, 33 (12) (2012) 1507–1512. DOI:10.1016/j.patrec.2012.04.013.
- [36] Mafikandi H., Amirkharian M., Solving Linear Partial Differential Equations by Moving Least Squares Method, *Bulletin of the Georgian National Academy of Sciences*, 9 (3) (2015) 26-36.
- [37] Mahalanobis P. Ch., On the generalized distance in statistics. *Proceedings of the National Institute of Sciences of India*, 2 (1) (1936) 49–55. Retrieved 2016-09-27.
- [38] Mao C., Hu B., Moore P., Su Y., Wang M. (2015) Nearest Neighbor Method Based on Local Distribution for Classification. In: Cao T., Lim EP., Zhou ZH., Ho TB., Cheung D., Motoda H. (eds) *Advances in Knowledge Discovery and Data Mining. PAKDD 2015. Lecture Notes in Computer Science*, vol 9077. Springer, Cham. DOI:10.1007/978-3-319-18038-0-19.
- [39] Melnykov I., Melnykov V., On K-means algorithm with the use of Mahalanobis distances, *Statistics & Probability Letters*, 84 (2014) 88-95. DOI:10.1016/j.spl.2013.09.026.
- [40] Michael R., O'Lenick C.R., Monaghan A., Wilhelm O., Wiedinmyer C., Hayden M., Estes M., Application of geostatistical approaches to predict the spatio-temporal distribution of summer ozone in Houston, Texas, *Journal of Exposure Science & Environmental Epidemiology* volume, 29 (2019) 806-820. DOI:10.1038/s41370-018-0091-4.
- [41] Monestiez P. , Courault I., Allard D., Ruget F., Spatial interpolation of air temperature using environmental context: Application to a crop model, *Environmental and Ecological Statistics*, 8 (2001) 297-309.
- [42] Ning X., Desrosiers C., Karypis G. (2015) A Comprehensive Survey of Neighborhood-Based Recommendation Methods. In: Ricci F., Rokach L., Shapira B. (eds) *Recommender Systems Handbook*. Springer, Boston, MA. DOI:10.1007/978-1-4899-7637-6-2.
- [43] Pauly M., Gross M., Kobbelt L. P., Efficient simplification of point-sampled surfaces, *IEEE Visualization. VIS 2002.*, Boston, MA, USA, (2002) 163-170. DOI: 10.1109/VISUAL.2002.1183771
- [44] Pauly M., Kobbelt L. P., Gross M., Point-based multiscale surface representation, *ACM Transactions on Graphics (TOG)*, 25 (2) (2006) 177–193. DOI: 10.1145/1138450.1138451.
- [45] Qasem SN, Samadianfar S, Sadri Nahand H, Mosavi A, Shamshirband S, Chau K-W. Estimating Daily Dew Point Temperature Using Machine Learning Algorithms. *Water*, 11 (3) (2019) 582.
- [46] Runions A., Samavati F., Prusinkiewicz P., Ribbons, A representation for point clouds. *The Visual Computer*, 23 (2007) 945–954. DOI:10.1007/s00371-007-0153-4.
- [47] Samworth R. J., Optimal weighted nearest neighbour classifiers. *Annals of Statistics*. 40 (5) (2012) 2733–2763. DOI:10.1214/12-AOS1049.

- 470 [48] Scofield B. C., A Signal from Saturn in Daily Temperature Data, Correlation, 29(1) 2013 53-66.
- [49] Sharma Sh., Mujumdar P.P., On the relationship of daily rainfall extremes and local mean temperature, Journal of Hydrology, 572 (2019) 179-191. DOI:10.1016/j.jhydrol.2019.02.048.
- [50] Shepard D., A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 1968 ACM National Conference. (1968) 517–524. DOI:10.1145/800186.810616.
- 475 [51] Stahl K., Moore R.D., Floyer J.A., Asplin M.G. , McKendry I.G., Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density, Agricultural and Forest Meteorology 139 (2006) 224–236. DOI:10.1016/j.agrformet.2006.07.004.
- [52] Stein M. S. (1999) Interpolation of Spatial Data; Some Theory for Kriging. New York, NY, USA: Springer-Verlag. DOI:10.1007/978-1-4612-1494-6.
- 480 [53] Viggiani M., Busetto L., Cimini D., Di Paola F., Geraldi E., Ranghetti L., Ricciardelli E., Romano F., A new spatial modeling and interpolation approach for high-resolution temperature maps combining reanalysis data and ground measurements, Agricultural and Forest Meteorology, (2019) 276–277, 107590. DOI:10.1016/j.agrformet.2019.05.021.
- [54] Wang B., Moving Least Squares Method for a One-Dimensional Parabolic Inverse Problem, Abstract and Applied Analysis, Volume 2014, Article ID 686020, 12 pages. DOI:10.1155/2014/686020.
- 485 [55] Wendland, H. (2004). Scattered Data Approximation (Cambridge Monographs on Applied and Computational Mathematics). Cambridge: Cambridge University Press. DOI:10.1017/CBO9780511617539.
- [56] You J., Hubbard K. G., Goddard S., Comparison of methods for spatially estimating station temperatures in a quality control system, International Journal of Climatology, 28 (2008) 777–787. DOI: 10.1002/joc.1571.
- 490 [57] Zhang GJ., Du JX., Huang DS., Lok TM., Lyu M.R. (2006) Adaptive Nearest Neighbor Classifier Based on Supervised Ellipsoid Clustering. In: Wang L., Jiao L., Shi G., Li X., Liu J. (eds) Fuzzy Systems and Knowledge Discovery. FSKD 2006. Lecture Notes in Computer Science, vol 4223. Springer, Berlin, Heidelberg. DOI:10.1007/11881599-69.